

MODÉLISATION,
ENRICHISSEMENT SÉMANTIQUE
ET DIFFUSION D'UN CORPUS
TEXTUEL SEMI-STRUCTURÉ: LE
CAS DES CATALOGUES DE VENTE
DE MANUSCRITS.

Paul, Hector Kervegan

25 septembre 2022

Introduction

Problématique

Introduction

Plan

- 1 La structure du texte comme méthode d'approche
 - Structure, document physique et encodage numérique
 - La spécificité d'un corpus "semi-structuré"
- 2 Sous quels angles aborder cette problématique?
 - Modéliser un corpus semi-structuré
 - Analyser le texte à partir de sa structure: la résolution d'entités nommées
 - Recomposer et diffuser le texte via une API
- 3 Le traitement automatique du texte comme chaîne éditoriale continue
 - L'enrichissement progressif du texte
 - La résolution d'entités nommées: mettre le texte en réseau
 - Recomposer le texte et créer des documents via une API

- 1 La structure du texte comme méthode d'approche
 - Structure, document physique et encodage numérique
 - La spécificité d'un corpus “semi-structuré”
- 2 Sous quels angles aborder cette problématique?
 - Modéliser un corpus semi-structuré
 - Analyser le texte à partir de sa structure: la résolution d'entités nommées
 - Recomposer et diffuser le texte via une API
- 3 Le traitement automatique du texte comme chaîne éditoriale continue
 - L'enrichissement progressif du texte
 - La résolution d'entités nommées: mettre le texte en réseau
 - Recomposer le texte et créer des documents via une API

La structure du texte comme méthode d'approche

Structure, document physique et encodage numérique

La structure du texte comme méthode d'approche
La spécificité d'un corpus "semi-structuré"

Plan

- 1 La structure du texte comme méthode d'approche
 - Structure, document physique et encodage numérique
 - La spécificité d'un corpus "semi-structuré"
- 2 Sous quels angles aborder cette problématique?
 - Modéliser un corpus semi-structuré
 - Analyser le texte à partir de sa structure: la résolution d'entités nommées
 - Recomposer et diffuser le texte via une API
- 3 Le traitement automatique du texte comme chaîne éditoriale continue
 - L'enrichissement progressif du texte
 - La résolution d'entités nommées: mettre le texte en réseau
 - Recomposer le texte et créer des documents via une API

Sous quels angles aborder cette problématique?
Modéliser un corpus semi-structuré

Sous quels angles aborder cette problématique?
Analyser le texte à partir de sa structure

Sous quels angles aborder cette problématique?

Recomposer et diffuser le texte via une API

Plan

- 1 La structure du texte comme méthode d'approche
 - Structure, document physique et encodage numérique
 - La spécificité d'un corpus "semi-structuré"
- 2 Sous quels angles aborder cette problématique?
 - Modéliser un corpus semi-structuré
 - Analyser le texte à partir de sa structure: la résolution d'entités nommées
 - Recomposer et diffuser le texte via une API
- 3 Le traitement automatique du texte comme chaîne éditoriale continue
 - L'enrichissement progressif du texte
 - La résolution d'entités nommées: mettre le texte en réseau
 - Recomposer le texte et créer des documents via une API

Une chaîne éditoriale continue
L'enrichissement progressif du texte

Une chaîne éditoriale continue

La résolution d'entités nommées: mettre le texte en réseau

Une chaîne éditoriale continue

Recomposer le texte et créer des documents via une API

Conclusion

There Is No Largest Prime Number

The proof uses *reductio ad absurdum*.

Theorem

There is no largest prime number.

- 1 Suppose p were the largest prime number.
- 2
- 3
- 4 But $q + 1$ is greater than 1, thus divisible by some prime number not in the first p numbers.

There Is No Largest Prime Number

The proof uses *reductio ad absurdum*.

Theorem

There is no largest prime number.

- 1 Suppose p were the largest prime number.
- 2 Let q be the product of the first p numbers.
- 3
- 4 But $q + 1$ is greater than 1, thus divisible by some prime number not in the first p numbers.

There Is No Largest Prime Number

The proof uses *reductio ad absurdum*.

Theorem

There is no largest prime number.

- 1 Suppose p were the largest prime number.
- 2 Let q be the product of the first p numbers.
- 3 Then $q + 1$ is not divisible by any of them.
- 4 But $q + 1$ is greater than 1, thus divisible by some prime number not in the first p numbers.